



1600

## RAW SEQUENCE LISTING

DATE: 07/17/2003

PATENT APPLICATION: US/09/868,546A

TIME: 15:37:49

Input Set : A:\BB-1324 Corrected Sequence Listing.txt

Output Set: N:\CRF4\07172003\I868546A.raw

3 <110> APPLICANT: E. I. du Pont de Nemours and Company  
 5 <120> TITLE OF INVENTION: Flavonoid Biosynthetic Enzymes  
 7 <130> FILE REFERENCE: BB1324  
 9 <140> CURRENT APPLICATION NUMBER: 09/868,546A  
 10 <141> CURRENT FILING DATE: 2001-09-20  
 12 <150> PRIOR APPLICATION NUMBER: 60/113,190  
 13 <151> PRIOR FILING DATE: 1998-12-21  
 15 <160> NUMBER OF SEQ ID NOS: 6  
 17 <170> SOFTWARE: Microsoft Office 97  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 1859  
 21 <212> TYPE: DNA  
 22 <213> ORGANISM: Glycine max  
 24 <400> SEQUENCE: 1

25	gaaaacactg	acagacagca	tagtctctg	tgcaagaatc	aattgagcaa	gcatgggaat	60
26	gttggtggtg	gtggtctcat	acgtgtctct	tttctggtt	ctattcctcg	gcgtgaagtt	120
27	tgttttccaa	agcagaaaat	tgagaaacat	accaccaggt	cctcctcctc	ttcccataat	180
28	aggaaacctt	aacctcctcg	aacagccaat	ccaccgtttc	ttccaacgca	tgtagaaaca	240
29	gtacggcaac	gtggtttccc	tctggttcgg	ttcacgtctg	gcggttgta	tctcctctcc	300
30	aacagcatac	caagaatgct	tcaccaaaca	cgacgttgcc	ttggccaacc	ggctaccttc	360
31	tctctcgga	aaatacatct	tctacaaca	caccaccgta	ggctcctgct	cccacggcga	420
32	gcactggcgc	aacctccgcc	gcatcacgc	cctggacgtc	ctctccacgc	agcgcgtcca	480
33	ctccttctcc	ggaatccgga	gcgacgagac	gaagcgtctg	atgcagaggt	tggtgctggc	540
34	caagaactcg	aacgaggaag	agtttgccgc	agtggagatt	agttcgatgt	tcaacgactt	600
35	aacttacaac	aacataatga	ggatgatata	ggggaagagg	ttttacggag	aggagagtga	660
36	gatgaagaac	gttgaggaag	cgagggagtt	cagagagact	gtgacagaaa	tggttgaact	720
37	catgggcttg	gctaacaagg	gagatcactt	gcctttcctc	aggtggttcg	attttcagaa	780
38	tgtggagaag	cgcttaaaaga	gtatcagtaa	gaggtacgat	tccatcttga	ataagatcct	840
39	tcatgagaac	cgtgccagca	atgaccgcca	gaattccatg	atcgatcata	tcctcaaaact	900
40	gcaagagacc	cagcctcagt	actacactga	ccaaatcata	aaaggccttg	ctctggccat	960
41	gctttttggt	ggaactgact	catcaactgg	gacttttagag	tggtcattat	ctaatttatt	1020
42	gaatcaccca	gaggtgttga	agaaggcaag	agatgaattg	gacactcaag	tgggacaaga	1080
43	ccgcttggtt	aatgagtcag	accttccaaa	acttccatat	cttaggaaga	tcatecctga	1140
44	gacacttagg	ttgtaccccc	cggccccaat	tctaatacct	catgtgtctt	cagaagatat	1200
45	tacaattgaa	ggattcaata	tcccacgaga	cacaattgtg	atcattaatg	gttggggcat	1260
46	gcagagagat	cctcagttgt	ggaatgatgc	cacatgcttt	aaacctgaga	ggtttgatgt	1320
47	ggaaggagag	gagaaaaagt	tggtagcatt	tggcatggga	agaagggctt	gccaggaga	1380
48	acccatggct	atgcaaagtg	tcagctttac	tttgggattg	ttgattcaat	gttttgactg	1440
49	gaaacgagta	agtgaggaaa	agcttgatat	gacagagaac	aattggatca	ccttgtcaag	1500
50	gttaattcca	ttggaggcca	tgtgcaaggc	tcgcccactt	gccactaaaa	ttggaattta	1560
51	attattaata	gtatttttat	ttggtaaact	tgggtgatcc	agaatctaata	acttataatt	1620
52	ttagtgtgta	agagtgggtga	tcatatatac	atttcaaaat	taataatctt	tgtccaaaaa	1680
53	tcatacatgg	acaactatat	gtcaattgac	atctagagag	aaatatagat	ataagaatat	1740

ENTERED

## RAW SEQUENCE LISTING

DATE: 07/17/2003

PATENT APPLICATION: US/09/868,546A

TIME: 15:37:49

Input Set : A:\BB-1324 Corrected Sequence Listing.txt

Output Set: N:\CRF4\07172003\I868546A.raw

```

54 ttatatttta ttactcttct ttatcttatg tgtcaaggcc cattgtagaa ttgggtgagc 1800
55 attaacatat atcaatattg tataccgccc agttttctca aataaatttc ttactttc 1859
57 <210> SEQ ID NO: 2
58 <211> LENGTH: 499
59 <212> TYPE: PRT
60 <213> ORGANISM: Glycine max
62 <400> SEQUENCE: 2
63 Leu Leu Val Val Val Ser Tyr Ala Val Leu Phe Leu Val Leu Phe Leu
64 1 5 10 15
66 Gly Val Lys Phe Val Phe Gln Ser Arg Lys Leu Arg Asn Ile Pro Pro
67 20 25 30
69 Gly Pro Pro Pro Leu Pro Ile Ile Gly Asn Leu Asn Leu Leu Glu Gln
70 35 40 45
72 Pro Ile His Arg Phe Phe Gln Arg Met Ser Lys Gln Tyr Gly Asn Val
73 50 55 60
75 Val Ser Leu Trp Phe Gly Ser Arg Leu Ala Val Val Ile Ser Ser Pro
76 65 70 75 80
78 Thr Ala Tyr Gln Glu Cys Phe Thr Lys His Asp Val Ala Leu Ala Asn
79 85 90 95
81 Arg Leu Pro Ser Leu Ser Gly Lys Tyr Ile Phe Tyr Asn Asn Thr Thr
82 100 105 110
84 Val Gly Ser Cys Ser His Gly Glu His Trp Arg Asn Leu Arg Arg Ile
85 115 120 125
87 Thr Ala Leu Asp Val Leu Ser Thr Gln Arg Val His Ser Phe Ser Gly
88 130 135 140
90 Ile Arg Ser Asp Glu Thr Lys Arg Leu Met Gln Arg Leu Val Leu Ala
91 145 150 155 160
93 Lys Asn Ser Asn Glu Glu Glu Phe Ala Arg Val Glu Ile Ser Ser Met
94 165 170 175
96 Phe Asn Asp Leu Thr Tyr Asn Asn Ile Met Arg Met Ile Ser Gly Lys
97 180 185 190
99 Arg Phe Tyr Gly Glu Glu Ser Glu Met Lys Asn Val Glu Glu Ala Arg
100 195 200 205
102 Glu Phe Arg Glu Thr Val Thr Glu Met Leu Glu Leu Met Gly Leu Ala
103 210 215 220
105 Asn Lys Gly Asp His Leu Pro Phe Leu Arg Trp Phe Asp Phe Gln Asn
106 225 230 235 240
108 Val Glu Lys Arg Leu Lys Ser Ile Ser Lys Arg Tyr Asp Ser Ile Leu
109 245 250 255
111 Asn Lys Ile Leu His Glu Asn Arg Ala Ser Asn Asp Arg Gln Asn Ser
112 260 265 270
114 Met Ile Asp His Leu Leu Lys Leu Gln Glu Thr Gln Pro Gln Tyr Tyr
115 275 280 285
117 Thr Asp Gln Ile Ile Lys Gly Leu Ala Leu Ala Met Leu Phe Gly Gly
118 290 295 300
120 Thr Asp Ser Ser Thr Gly Thr Leu Glu Trp Ser Leu Ser Asn Leu Leu
121 305 310 315 320
123 Asn His Pro Glu Val Leu Lys Lys Ala Arg Asp Glu Leu Asp Thr Gln
124 325 330 335

```

## RAW SEQUENCE LISTING

DATE: 07/17/2003

PATENT APPLICATION: US/09/868,546A

TIME: 15:37:49

Input Set : A:\BB-1324 Corrected Sequence Listing.txt

Output Set: N:\CRF4\07172003\I868546A.raw

```

126 Val Gly Gln Asp Arg Leu Leu Asn Glu Ser Asp Leu Pro Lys Leu Pro
127                               340                               345                               350
129 Tyr Leu Arg Lys Ile Ile Leu Glu Thr Leu Arg Leu Tyr Pro Pro Ala
130                               355                               360                               365
132 Pro Ile Leu Ile Pro His Val Ser Ser Glu Asp Ile Thr Ile Glu Gly
133                               370                               375                               380
135 Phe Asn Ile Pro Arg Asp Thr Ile Val Ile Ile Asn Gly Trp Gly Met
136 385                               390                               395                               400
138 Gln Arg Asp Pro Gln Leu Trp Asn Asp Ala Thr Cys Phe Lys Pro Glu
139                               405                               410                               415
141 Arg Phe Asp Val Glu Gly Glu Glu Lys Lys Leu Val Ala Phe Gly Met
142                               420                               425                               430
144 Gly Arg Arg Ala Cys Pro Gly Glu Pro Met Ala Met Gln Ser Val Ser
145                               435                               440                               445
147 Phe Thr Leu Gly Leu Leu Ile Gln Cys Phe Asp Trp Lys Arg Val Ser
148                               450                               455                               460
150 Glu Glu Lys Leu Asp Met Thr Glu Asn Asn Trp Ile Thr Leu Ser Arg
151 465                               470                               475                               480
153 Leu Ile Pro Leu Glu Ala Met Cys Lys Ala Arg Pro Leu Ala Thr Lys
154                               485                               490                               495
156 Ile Gly Ile
159 <210> SEQ ID NO: 3
160 <211> LENGTH: 1698
161 <212> TYPE: DNA
162 <213> ORGANISM: Glycine max
164 <400> SEQUENCE: 3
165 cagtaataac aatgtctcct ttcttatctt actctcttct ttcctcgtg ttctttctca 60
166 ctctcaagta ccttttccaa agaagcagaa aagtacgaaa cctgccacct ggtccgactc 120
167 ctcttccctat aatcggaac cttaacctcg ttgaacaacc tatacacctg ttcttccacc 180
168 gcatgtccca aaaatatgga aacatcatat ccttttggtt tgggtcacgt cttgttgtgg 240
169 ttgtttcatc acccacagcg taccaagaat gtttcaccaa acatgatgtt accttgacca 300
170 acagggtagc ctccctctcg ggaaaataca tattctacga caacaccacc gtagggctct 360
171 gctcccacgg cgagcactgg cgcaacctcc gccgcataac ctctctcgac gttctatcga 420
172 cgcagcgcgt ccactccttc tccggaatcc ggagcgacga gacgaagagg ttgatacaca 480
173 ggctggccag ggactccggg aaagattttg cgcgcgtgga gatgacctcc aagtttgctg 540
174 acttgacgta caacaacatc atgaggatga tttcggggaa gcggttttac ggagaagaga 600
175 gtgaacttaa caacgttgag gaagcgaagg agttcagaga cactgtgaat gagatgctgc 660
176 aactcatggg gttggctaac aaggggagatc acttaccttt cctaagggtg ttcgattttc 720
177 agaacgtgga gaagagggtg aagaatatca gtaagaggta tgataccatc ttgaataaga 780
178 tccttgatga gaaccgtaac aacaaggacc gcgagaattc catgattggt catctcctca 840
179 aactgcaaga gacacagcct gactattata ccgaccaa atcatcaaagg cttgcttttg 900
180 ctatgtctct tgggtggaaca gactcgtcaa ctggaacttt agagtgggca ttatctaatt 960
181 tagtgaatga cccagagggt ctgcagaagg caagagatga gttggacgct caagtaggac 1020
182 catatcggtt gttaaatgag tcagaccttc caaaacttcc ttatctcagg aagatagttc 1080
183 ttgaaacact taggttgtac cctccggctc caattcta ataccacgtg gcttcagaag 1140
184 acatcaatat cgaaggattc aatgttccac gagacaca attgtgattatt aatggttggg 1200
185 ccatgcaaag agatcctaag atatggaaag atgcgacaag ctttaaacct gagaggtttg 1260
186 atgaagaagg agaggagaag aaattggtag catttggtat gggagaagg gcttgccag 1320
187 gagaacccat ggctatgcaa agtgtttagct atactttggg attaattgatt caatgttttg 1380

```

## RAW SEQUENCE LISTING

DATE: 07/17/2003

PATENT APPLICATION: US/09/868,546A

TIME: 15:37:49

Input Set : A:\BB-1324 Corrected Sequence Listing.txt

Output Set: N:\CRF4\07172003\I868546A.raw

```

188 actggaaacg agtaagtgag aagaagcttg atatgacaga gaataattgg atcaccttgt 1440
189 caagggttaat tccattggag gctatgtgta aagcccgccc actcgccagc aaagttgaaa 1500
190 gttattaaca atattttatt tggatatatt gggtgaggat ctaataactca taatttcggt 1560
191 gtgtaagtct atgcatgtta aaattaataa tatttgctgt atgtccacaa ggccaaatgt 1620
192 agtactgggt gtggatttgc atatacaata tcaatattgt ataaatccca gtttccttga 1680
193 ataaatttct ttactttc                                     1698
195 <210> SEQ ID NO: 4
196 <211> LENGTH: 494
197 <212> TYPE: PRT
198 <213> ORGANISM: Glycine max
200 <400> SEQUENCE: 4
201 Leu Ser Tyr Ser Leu Leu Ser Leu Val Phe Phe Phe Thr Leu Lys Tyr
202   1           5           10           15
204 Leu Phe Gln Arg Ser Arg Lys Val Arg Asn Leu Pro Pro Gly Pro Thr
205           20           25           30
207 Pro Leu Pro Ile Ile Gly Asn Leu Asn Leu Val Glu Gln Pro Ile His
208           35           40           45
210 Arg Phe Phe His Arg Met Ser Gln Lys Tyr Gly Asn Ile Ile Ser Leu
211           50           55           60
213 Trp Phe Gly Ser Arg Leu Val Val Val Val Ser Ser Pro Thr Ala Tyr
214   65           70           75           80
216 Gln Glu Cys Phe Thr Lys His Asp Val Thr Leu Ala Asn Arg Val Arg
217           85           90           95
219 Ser Leu Ser Gly Lys Tyr Ile Phe Tyr Asp Asn Thr Thr Val Gly Ser
220           100          105          110
222 Cys Ser His Gly Glu His Trp Arg Asn Leu Arg Arg Ile Thr Ser Leu
223           115          120          125
225 Asp Val Leu Ser Thr Gln Arg Val His Ser Phe Ser Gly Ile Arg Ser
226           130          135          140
228 Asp Glu Thr Lys Arg Leu Ile His Arg Leu Ala Arg Asp Ser Gly Lys
229 145           150          155          160
231 Asp Phe Ala Arg Val Glu Met Thr Ser Lys Phe Ala Asp Leu Thr Tyr
232           165          170          175
234 Asn Asn Ile Met Arg Met Ile Ser Gly Lys Arg Phe Tyr Gly Glu Glu
235           180          185          190
237 Ser Glu Leu Asn Asn Val Glu Glu Ala Lys Glu Phe Arg Asp Thr Val
238           195          200          205
240 Asn Glu Met Leu Gln Leu Met Gly Leu Ala Asn Lys Gly Asp His Leu
241           210          215          220
243 Pro Phe Leu Arg Trp Phe Asp Phe Gln Asn Val Glu Lys Arg Leu Lys
244 225           230          235          240
246 Asn Ile Ser Lys Arg Tyr Asp Thr Ile Leu Asn Lys Ile Leu Asp Glu
247           245          250          255
249 Asn Arg Asn Asn Lys Asp Arg Glu Asn Ser Met Ile Gly His Leu Leu
250           260          265          270
252 Lys Leu Gln Glu Thr Gln Pro Asp Tyr Tyr Thr Asp Gln Ile Ile Lys
253           275          280          285
255 Gly Leu Ala Leu Ala Met Leu Phe Gly Gly Thr Asp Ser Ser Thr Gly
256           290          295          300

```

## RAW SEQUENCE LISTING

DATE: 07/17/2003

PATENT APPLICATION: US/09/868,546A

TIME: 15:37:49

Input Set : A:\BB-1324 Corrected Sequence Listing.txt

Output Set: N:\CRF4\07172003\I868546A.raw

```

258 Thr Leu Glu Trp Ala Leu Ser Asn Leu Val Asn Asp Pro Glu Val Leu
259 305 310 315 320
261 Gln Lys Ala Arg Asp Glu Leu Asp Ala Gln Val Gly Pro Asp Arg Leu
262 325 330 335
264 Leu Asn Glu Ser Asp Leu Pro Lys Leu Pro Tyr Leu Arg Lys Ile Val
265 340 345 350
267 Leu Glu Thr Leu Arg Leu Tyr Pro Pro Ala Pro Ile Leu Ile Pro His
268 355 360 365
270 Val Ala Ser Glu Asp Ile Asn Ile Glu Gly Phe Asn Val Pro Arg Asp
271 370 375 380
273 Thr Ile Val Ile Ile Asn Gly Trp Ala Met Gln Arg Asp Pro Lys Ile
274 385 390 395 400
276 Trp Lys Asp Ala Thr Ser Phe Lys Pro Glu Arg Phe Asp Glu Glu Gly
277 405 410 415
279 Glu Glu Lys Lys Leu Val Ala Phe Gly Met Gly Arg Arg Ala Cys Pro
280 420 425 430
282 Gly Glu Pro Met Ala Met Gln Ser Val Ser Tyr Thr Leu Gly Leu Met
283 435 440 445
285 Ile Gln Cys Phe Asp Trp Lys Arg Val Ser Glu Lys Lys Leu Asp Met
286 450 455 460
288 Thr Glu Asn Asn Trp Ile Thr Leu Ser Arg Leu Ile Pro Leu Glu Ala
289 465 470 475 480
291 Met Cys Lys Ala Arg Pro Leu Ala Ser Lys Val Glu Ser Tyr
292 485 490
294 <210> SEQ ID NO: 5
295 <211> LENGTH: 843
296 <212> TYPE: DNA
297 <213> ORGANISM: Glycine max
299 <220> FEATURE:
300 <221> NAME/KEY: unsure
301 <222> LOCATION: (476)
302 <223> OTHER INFORMATION: n = a, c, g or t
304 <220> FEATURE:
305 <221> NAME/KEY: unsure
306 <222> LOCATION: (657)
307 <223> OTHER INFORMATION: n = a, c, g or t
309 <220> FEATURE:
310 <221> NAME/KEY: unsure
311 <222> LOCATION: (703)
312 <223> OTHER INFORMATION: n = a, c, g or t
314 <220> FEATURE:
315 <221> NAME/KEY: unsure
316 <222> LOCATION: (712)
317 <223> OTHER INFORMATION: n = a, c, g or t
319 <220> FEATURE:
320 <221> NAME/KEY: unsure
321 <222> LOCATION: (789)
322 <223> OTHER INFORMATION: n = a, c, g or t
324 <220> FEATURE:

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 07/17/2003  
PATENT APPLICATION: US/09/868,546A      TIME: 15:37:50

Input Set : A:\BB-1324 Corrected Sequence Listing.txt  
Output Set: N:\CRF4\07172003\I868546A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 476,657,703,712,789,843